

Calculating Gini

Weekend	Weather	Parents	Money	Decision
W ₁	Sunny	Yes	Rich	Cinema.
W ₂	Sunny	No	Rich	Tennis
W ₃	Windy	Yes	Rich	Cinema.
W ₄	Rainy	Yes	Poor	Cinema
W ₅	Rainy	No	Rich	Stay In
W ₆	Rainy	Yes	Poor	Cinema.
W ₇	Windy	No	Poor	Cinema.
W ₈	Windy	No	Rich	Shopping
W ₉	Windy	Yes	Rich	Cinema.
W ₁₀	Sunny	No	Rich	Tennis

Attributes $\rightarrow 3$

Output variable \rightarrow Decisions.

Possible outputs \rightarrow Cinema, Tennis, Stay In, Shopping.

Decision Tree ?

1st: Find out the ~~maxim~~ attribute having maximum information gain or least gini.

Step I : Gini for the total data set.

Possible outputs = 4

No of Instances = 10.

\therefore Cinema $\rightarrow 6$

Tennis $\rightarrow 2$

Stay In $\rightarrow 1$

Shopping $\rightarrow 1$

$$\begin{aligned} \text{Gini} &= 1 - \left[\left(\frac{6}{10}\right)^2 + \left(\frac{2}{10}\right)^2 + \left(\frac{1}{10}\right)^2 + \left(\frac{1}{10}\right)^2 \right] \\ &= 1 - [0.36 + 0.04 + 0.0001] \\ &= \underline{0.5999} \end{aligned}$$

Gini Index for Attribute = Money

Instances; 1. Rich = 7

2. Poor = 3

For, Money = Rich * Out of 7 instances;

Tennis = 2

Cinema = 3

Shopping = 1

Stay in = 1

$$\begin{aligned} \therefore \text{Gini} &= 1 - \left[\left(\frac{2}{7}\right)^2 + \left(\frac{3}{7}\right)^2 + \left(\frac{1}{7}\right)^2 + \left(\frac{1}{7}\right)^2 \right] \\ &= \underline{0.6938} \end{aligned}$$

Money = Poor

Cinema = 3

$$\begin{aligned} \text{Gini} &= 1 - \left[\left(\frac{3}{3}\right)^2 \right] \\ &= \underline{0} \end{aligned}$$

Weighted Average / Gini for attribute = Money

$$\begin{aligned} &= \left(0 \times \frac{3}{10}\right) + \left(0.6938 \times \frac{7}{10}\right) \\ &= \underline{0.48566} \end{aligned}$$

* Gini for Parents

Yes = 5

No = 5

Parents = Yes ; Cinema = 5 (All 5 outputs are related to Cinema)

Parents = No ; Tennis = 2 Cinema = 1

Stay in = 1

Shopping = 1

$$\therefore \text{Gini} = 1 - \left[\left(\frac{2}{5}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(\frac{1}{5}\right)^2 \right]$$

$$= \underline{0.72}$$

$$\text{Gini}_{(\text{Parents})} = \left(0 \times \frac{5}{10} \right) + \left[0.72 \times \left(\frac{5}{10}\right) \right]$$

$$= \underline{0.36}$$

\therefore For weather

$$\text{Gini}_{(\text{sunny})} = 1 - \left[\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 \right]$$

$$= \underline{0.444}$$

$$\text{Gini}_{(\text{Rainy})} = 1 - \left[\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 \right]$$

$$= \underline{0.444}$$

$$\text{Gini}_{(\text{windy})} = 1 - \left[\left(\frac{3}{4}\right)^2 + \left(\frac{1}{4}\right)^2 \right]$$

$$= \underline{0.375}$$

$$\text{Gini}_{(\text{weather})} = 0.444 \times \left(\frac{3}{10}\right) + 0.444 \times \left(\frac{3}{10}\right) + 0.375 \times \left(\frac{4}{10}\right)$$

$$= \underline{0.416}$$

<u>Gini Index</u>	<u>Value</u>
-------------------	--------------

Weather	0.416
---------	-------

Parents	0.36
---------	------

Money	0.486
-------	-------

→ Less

→ Select Parent as the 'Root Node'.

→ Divide the data into possible values of 'Parent' i.e., Yes or No.

Parent

Decision = Cinema.

Yes

No

W	Weather	Parents	Money	Decision
W ₁	Sunny	Yes	Rich	Cinema
W ₂	Windy	Yes	Rich	Cinema
W ₃	Rainy	Yes	Poor	Cinema
W ₄	Rainy	Yes	Poor	Cinema
W ₅	Windy	Yes	Rich	Cinema

W	Weather	Parents	Money	Decision
W ₂	Sunny	No	Rich	Tennis
W ₅	Rainy	No	Rich	Stay In
W ₇	Windy	No	Poor	Cinema
W ₈	Windy	No	Rich	Shopping
W ₁₀	Sunny	No	Poor	Tennis

If (Parent = Yes, whether (Weather = Sunny or Windy or Rainy, Money = Rich or Poor))
The decision is → Cinema.

Parent = No, Attribute = Weather.

Sunny = 2

Windy = 2

Rainy = 1

Tennis = 2.

Rainy = 1

Gini (Weather = Sunny)

$$= 1 - \left(\frac{2}{2}\right)^2$$

$$= \underline{\underline{0}}$$

Weather = Windy

Cinema = 1

Shopping = 1

∴ Gini (Weather = Windy)

$$= 1 - \left[\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 \right]$$

$$= \underline{\underline{0.5}}$$

Weather = Rainy.

Stay In = 1

∴ Gini (Weather = Rainy)

$$= 1 - \left(\frac{1}{1}\right)^2$$

$$= \underline{\underline{0}}$$

Weighted Average :

$$= 0 \times \frac{2}{5} + 0 \times \frac{1}{5} + 0.5 \times \left(\frac{2}{5}\right)$$

$$= \underline{\underline{0.2}}$$

